

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-251572

(43) 公開日 平成8年(1996)9月27日

(51) Int.Cl. <sup>8</sup>	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 N 7/173			H 0 4 N 7/173	
H 0 4 M 11/06			H 0 4 M 11/06	
H 0 4 N 1/00	1 0 2		H 0 4 N 1/00	1 0 2 A
5/93			5/93	E

審査請求 未請求 請求項の数 9 F D (全 7 頁)

(21) 出願番号 特願平7-329480

(22) 出願日 平成7年(1995)11月27日

(31) 優先権主張番号 3 4 5 6 7 8

(32) 優先日 1994年11月28日

(33) 優先権主張国 米国 (US)

(71) 出願人 390035493

エイ・ティ・アンド・ティ・コーポレーション

AT&amp;T CORP.

アメリカ合衆国 10013-2412 ニューヨーク  
ニューヨーク アヴェニュー オブ  
ジ アメリカズ 32

(72) 発明者 ジョン ダブリュー・コブル

アメリカ合衆国, 43230 オハイオ, ガハ  
ンナ, チルトン プレイス 360

(74) 代理人 弁理士 三俣 弘文

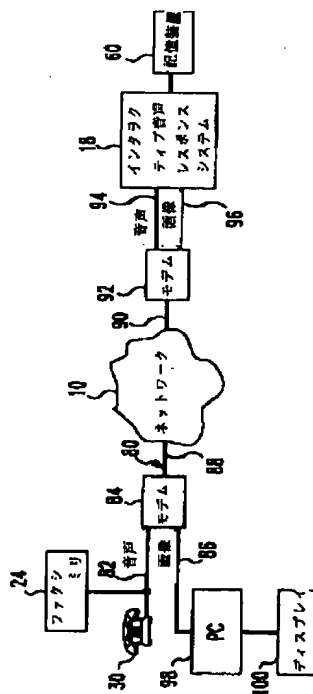
最終頁に続く

(54) 【発明の名称】 画像データ伝送方法

(57) 【要約】

【課題】 顧客が音声コマンドを用いて情報通信ネットワーク10で画像および/または音声の情報を伝送する情報伝送を制御する方法を提供する。

【解決手段】 発呼者から発話をネットワーク10で受信するステップと、音声認識装置18を用いてこの発話を翻訳するステップと、この翻訳ステップに応答し記憶装置60から画像データを検索するステップと、この画像データをネットワーク10で送信するステップを有する方法を提供する。尚この画像データはファクシミリ画像、ビデオ画像または動画ビデオ画像を示す。



## 【特許請求の範囲】

【請求項 1】 (A) 発呼者から発話を情報通信ネットワーク (10) で受信する受信ステップと、

(B) 音声認識装置 (18) を用いて前記発話を翻訳する翻訳ステップと、

(C) 前記発話を翻訳する前記翻訳ステップに応答し記憶装置 (60) から画像データを検索する検索ステップと、

(D) 前記画像データをこの情報通信ネットワーク (10) で送信する送信ステップを有することを特徴とする情報通信ネットワーク (10) で画像データを伝送する画像データ伝送方法。

【請求項 2】 前記画像データはファクシミリを表すことを特徴とする請求項 1 に記載の方法。

【請求項 3】 前記発話をこの情報通信ネットワーク 10 の第 1 の通信チャネルで通信しおよび前記画像データをこの情報通信ネットワークの第 2 の通信チャネルで通信することを特徴とする請求項 1 に記載の方法。

【請求項 4】 前記発話と前記画像データをこの情報通信ネットワーク 10 の単一通信チャネルで通信することを特徴とする請求項 1 に記載の方法。

【請求項 5】 前記検索ステップは音声データを検索するステップを有し、および前記送信ステップは音声データを送信するステップを有することを特徴とする請求項 1 に記載の方法。

【請求項 6】 前記画像と前記音声のデータを同時に送信することを特徴とする請求項 5 に記載の方法。

【請求項 7】 前記発話をこの情報通信ネットワーク 10 の第 1 の通信チャネルで通信し、前記画像と前記音声のデータをこの情報通信ネットワーク 10 の第 2 の通信チャネルで通信することを特徴とする請求項 6 に記載の方法。

【請求項 8】 前記画像と前記音声のデータを時分割多重化を用いて送信することを特徴とする請求項 5 に記載の方法。

【請求項 9】 前記発話をこの情報通信ネットワーク 10 の第 1 の通信チャネルで通信し、前記画像と前記音声のデータをこの情報通信ネットワーク 10 の第 2 の通信チャネルで通信することを特徴とする請求項 8 に記載の方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、情報通信ネットワークで画像および／または音声の情報を伝送する方法に係り、特に音声コマンドを用いて情報通信ネットワークで画像および／または音声の情報を伝送する情報伝達を顧客が制御する方法に関する。

## 【0002】

【従来の技術】従来、ホーム・ショッピングのようなサービスは、その顧客の住宅へのカタログ郵便やホーム・

ショッピングのテレビジョン・プログラムで視聴して注文を電話で行った。カタログ・ショッピングでは顧客がその顧客の好都合な際に商品の絵を見て行うことができるが、新しいカタログがこの顧客に郵送されるまでそのカタログは提供される最新商品でカタログが更新されることはできない。この顧客にホーム・ショッピング・テレビジョン・プログラムは提供商品の最新の情報を供与する。しかしこの顧客は無理に特定のプログラムのスケジュールの際にショッピングすることになる。さらにこの顧客はそのテレビジョン画面上の提供商品画像を制御できない。そのためこの顧客は関心のない商品にも無理に視聴する結果となる。次のように以上の難点を解消する方法が望まれている。

## 【0003】

【発明が解決しようとする課題】音声コマンドを用いて情報通信ネットワークで画像および／または音声の情報を伝送する情報伝達を顧客が制御する方法が望まれている。

## 【0004】

【課題を解決するための手段】本発明は以下に説明する方法を顧客に提供して前記課題を解決しこの技術分野の進歩を遂げる。すなわち、顧客がその情報通信ネットワークを利用しその情報通信ネットワークでこの顧客に伝達される商品画像を選択することができる。つまり、本発明の方法によってこの顧客に関心のある商品画像を選択しながらこの顧客の都合に応じこの顧客はショッピングすることができる。この顧客は音声コマンドを用いて提供商品画像を制御し購入商品を選択できる。

## 【0005】

【発明の実施の形態】図 1 に情報通信ネットワーク 10 を示しこれには通信チャネル 12、14、16 がある。通信チャネル 12 は、コンバーサント (R) (Conversant (R)) (本例では商標登録済表示の丸 R 印を (R) と表記する) インタラクティブ音声レスポンス (IVR) システム 18 (ただし Conversant (R) は A T & T 社の登録商標である) から信号バス 20 で情報通信ネットワーク 10 を経てさらに信号バス 22 でファクシミリ端末 24 へファクシミリ画像データを伝送する (コンバーサント (R) IVR システムは A T & T 社から提供され利用できる)。通信チャネル 14 を用いて、電話機器 30 から信号バス 32 で情報通信ネットワーク 10 を経て次に信号バス 34 でコンバーサント IVR システム 18 へ音声通信を提供する。

## 【0006】通信チャネル 16 を用いて、ピカソ (R)

(Picasso (R)) 電話 40 (ただし Picasso (R) は A T & T 社の登録商標である) から信号バス 42 で情報通信ネットワーク 10 を経て信号バス 44 でピカソ電話 46 へビデオ画像データを提供する。ピカソ電話 46 は信号バス 16 から受信したビデオ画像をテレビジョンまたはビデオ・モニタ 48 上に表示する (ピカソ電話は A T &

T社から入手できる)。顧客は電話機器30と通信チャンネル14を用いてコンバーサントIVRシステム18と接触する。コンバーサントIVRシステム18は自動音声認識機能を用いて通信チャンネル14でこの顧客によって提供されたコマンドを翻訳する。

【0007】このコンバーサントIVRシステムは通信チャンネル14でこの顧客にオプションのメニューを提供し、この顧客は通信チャンネル14で音声コマンドを用いてこれらのメニューから選択を行う。これらのメニューによって、例えば、顧客は画像が送信される前に、数種類の商品から選択を行うことができる。このようにこの顧客はこの顧客の関心のある商品の画像のみを要求することができる。他の例では、顧客は、遠隔からこのシステムの利用によって不動産を調査することができる。この顧客は、調査しようとする不動産に対し特定の都市や価格範囲を選択するためオプションのメニューに口頭で応答することができる。次にこのコンバーサントIVRシステムは通信チャンネル16でこの顧客の要求を満足させる不動産の種々の物件のビデオ画像を送信する。

【0008】この顧客はこの顧客の関心のある不動産物件を見付けると、この顧客は音声コマンドを用いてその不動産物件内の例えば、部屋のようなビデオ画像の詳細追加を受信することができる。これらの画像には、例えば、台所、リビング・ルーム、地下室などのようなビデオ画像がある。さらにまたこの顧客はデータ表によって商品や不動産物件の追加情報を受信することができるが、ただしこれは通信チャンネル12でそのコンバーサントIVRシステムからこのユーザのファクシミリ端末へファックスで送付される。前述のように、ユーザは電話機器30を用いて通信チャンネル14でコンバーサントIVRシステム18に最初に接触する。

【0009】コンバーサントIVRシステム18はそのユーザに選択メニューを提供し音声認識装置を用いてその顧客の選択に応答する。このコンバーサントIVRシステムは内部メモリを持ちこれがスクリプトを提供しこれを用いてその顧客にメニューを提供することができ、または信号バス62でコンバーサントIVRシステム18と通信する外部記憶装置60を用いることができる。記憶装置60には、例えば、ハード・ディスク、光ディスク、CDロム、テープ・ユニット、または他の種類のストレージ、例えば、メモリであってホスト・コンピュータまたはサーバ上のものを挙げることができる。信号バス62にはインタフェースがあつてこれには、例えば、SCSIインタフェース、標準のバス・インタフェース、LANインタフェースおよび/または情報通信ネットワークへのインタフェースを挙げることができる。

【0010】このコンバーサントIVRシステムがその顧客に選択メニューを提供しさらに音声認識機能を利用してこの顧客の音声コマンドまたは発話を翻訳した後、コンバーサントIVRシステム18はその選択されたビデ

オ画像をその顧客に通信チャンネル16で提供する。さらにコンバーサントIVRシステム18が一連の画像を提供することも可能であつてこれによってこの顧客はテレビジョンまたはモニタ48で動画ビデオを視聴することができる。このコンバーサントIVRシステムは内部記憶装置または記憶装置60のようなストレージを用いてそのビデオ画像を供与しその選択画像を表すデータをアクセスできる。この画像は、ISO 10918-1/CCITT JPEG標準フォーマット(JPEG)のような標準フォーマットでデジタル形で通常蓄積される。

【0011】次にこのコンバーサントIVRシステムは、ピカソ電話40へ並列通信バス64でJPEGフォーマットでこのビデオ画像データを提供する。次にこのピカソ電話はコンバーサントIVRシステム18によって提供された画像データを信号バス42で情報通信ネットワーク10を経てさらに信号バス44で受信ピカソ電話46へ転送する。ピカソ電話46は、NTSPビデオ信号のようなビデオ信号をテレビジョンまたはモニタ48に提供してこのJPEGデータ・ファイルをビデオ画像に変換する。さらにまた音声情報をそのビデオ画像データと共に含むことも可能である。この場合には、コンバーサントIVRシステム18はこの音声データをそれ自身のメモリまたは記憶装置60のような記憶装置から得る。

【0012】この音声情報を通信バス34で電話機器30へ伝える。さらにまたこの音声情報をピカソ電話40を用いて送ることも可能である。ピカソ電話40はこのJPEGビデオ画像データとその音声データを通信チャンネル16で受信ピカソ電話46へ送信する。次にピカソ電話46はこの音声とビデオ画像のデータを分離してこのビデオ画像データがテレビジョン48で画像を生成しまたこの音声データをこのピカソ電話に含まれるスピーカフォンによって再生する。このピカソ電話以外の装置も、この送信装置とその受信装置がこのビデオ画像および/またはその音声のデータをパッケージまたは伝送するために同じまたはコンパチブルなフォーマットを使用する限り、それを利用することができる。

【0013】ピカソ電話の場合には、その画像および/または音声のデータはX.25フォーマット(X.25はCCITT標準である)で伝送する。この顧客が通信チャンネル14を用いてファクシミリ画像データを要求する場合、コンバーサントIVRシステム18は、内部メモリまたは記憶装置60のような記憶装置からその蓄積ファクシミリ画像データをアクセスしこのファクシミリ画像データを標準ファクシミリ・フォーマットで信号バス20で情報通信ネットワーク10を経てさらに信号バス22でファクシミリ端末24に転送する。このコンバーサントIVRシステムによって利用されるスクリプトは、AT&T社から購入またはAT&T社から入手できるスクリプト・ビルディング・ソフトウェア・ツールを

用いる開発/特注で入手することができる。

【0014】音声コマンドに加えて、この顧客は電話機器30でキーボード入力によってコンバーサントIVRシステム18にコマンドを伝達できる。キーボード入力の場合にはコンバーサントIVRシステム18は顧客コマンドを翻訳するため電話機器30が提供するDTMF信号をデコードする。図2に本発明の別の実施の形態例を示す。この実施の形態の提供する利点は単一の通信チャネルを利用してこの顧客とそのコンバーサントIVRシステム間のすべての通信を行う点である。つまりこの実施の形態ではこの顧客の発話ならびにコンバーサントIVRシステム18が供給するビデオ画像および/または音声のデータは単一の通信チャネルで情報通信ネットワーク10を経て伝えられる。本例では通信チャネル80を用いて、この顧客から音声コマンドまたは発話をコンバーサントIVRシステム18へ提供しさらにコンバーサントIVRシステム18からこの顧客へビデオ画像と音声の情報を提供する。

【0015】この顧客は、音声コマンドを電話機器30を経て信号バス82でボイスパンTMモデム84 (VoiceSpan はAT&T社の商標である) (ボイスパンTMモデムはAT&T社から提供され利用できる) に入力する。モデム84は、信号バス82で電話機器30から音声情報のようなアナログ情報またはデータを、信号バス86でモデム84に提供されたデジタル・データまたは情報と結合する。(本例ではモデム84が信号バス86でデジタル情報を受信しない。しかしクレジット・カードの画像または口座番号を表すデータのようなデジタル・データをコンピュータ98と信号バス86を介してモデム84に提供できる。) モデム84はこのユーザの発話を信号バス88で情報通信ネットワーク10を経て信号バス90へ送信する。

【0016】ボイスパンTMモデム92は信号バス90で情報通信ネットワーク10からその信号を受信しこれをアナログとデジタルの出力に分離する。アナログ出力は信号バス94でコンバーサントIVRシステム18へまたデジタル出力は信号バス96でコンバーサントIVRシステム18へ提供される。(前述のように、本例ではモデム84からモデム92に進むデジタル情報はない。また前述のように、PC98がモデム84、92を用いてクレジット・カード情報のような情報をコンバーサントIVRシステム18へ伝えることができる。) コンバーサントIVRシステム18はこの音声メニューを電話機器30におけるそのユーザへ信号バス94とモデム92を介しチャネル80を通りモデム84と信号バス82を経て伝える。

【0017】次ぎにこのユーザはこのメニューから選択するよう音声コマンドを入力する。この音声コマンドはコンバーサントIVRシステム18によって翻訳されまたこれを用いてこのユーザに追加のメニューまたはその選択

画像を提供する。先に図1に関して述べたように、選択画像に音声情報を含むことができる。さらに一連のビデオ画像を提供しこの顧客が動画ビデオを受信することができる。コンバーサントIVRシステム18は、この顧客が選択したビデオ画像データおよび/または音声データを提供するがこれはコンバーサントIVRシステム18のメモリまたは記憶装置60内に蓄積した音声ファイルおよび/またはビデオ画像ファイルをアクセスして行われる。

10 【0018】この音声データはアナログ形で信号バス94を介しモデム92に提供されさらにこのビデオ画像データは信号バス96を介しモデム92に提供される。前述のように、この画像データはJPEGフォーマットのようないずれかの通常使用されるフォーマットとすることができる。モデム92はこのアナログ音声信号とそのデジタル・ビデオ画像データを結合しそして同時にこの音声と画像のデータを信号バス90で情報通信ネットワーク10を経て信号バス80へさらにモデム84へ送る。モデム84はこの音声と画像のデータを分離しこの音声信号を通信バス82で電話機器30へこの顧客の聴取用に提供する。モデム84はこのビデオ画像データを信号バス86でパーソナル・コンピュータ98に提供する。

20 【0019】パーソナル・コンピュータ98はモデム84から受信した標準のビデオ画像データ・ファイルを用いてこの送信ビデオ画像をディスプレイまたはモニタ100に生成する。ここで注記する必要があることは、この顧客にJPEGタイプの画像データを送信することに加えて、さらにファクシミリ画像データをこの顧客に信号バス94かまたは信号バス96かのいずれかで送信することができる。このファクシミリ画像データが信号バス96で送られるとすると、モデム84はこの情報をPC98に信号バス86を通して提供する。そこでPC98はこのファクシミリ画像データをディスプレイ100に表示またはプリンタを用いてこのファクシミリ画像をプリントする。

30 【0020】このファクシミリ画像データがアナログ形で信号バス94を通して送られると、モデム84はこの出力ファクシミリ画像データをアナログ形で信号バス82に提供しファクシミリ端末24がこのファクシミリ画像を生成することができる。ここでまた注記する必要があることは、ボイスパンTMモデムはこのアナログ音声データとデジタル画像データを同時に送信できることである。例えば、モデム92は信号バス94からこの音声情報と信号バス96からその画像データを同時に送信しその一方でモデム84は音声情報と画像データを信号バス88から同時に受信する。さらにまたモデムまたは同様の装置であって時分割多重化または周波数分割多重化を用いる装置を利用してこの音声と画像のデータを送ることができる。

【0021】図3に本発明の別の実施の形態例を示す。この実施の形態例の提供する利点もまた単一通信チャネルを用いてこの顧客とそのコンバーサントIVRシステム間のすべての通信を提供することのある。この実施の形態例ではこの顧客の発話とコンバーサントIVRシステム18が供給するビデオ画像および／または音声のデータは両者とも単一通信チャネルで情報通信ネットワーク10を経て伝達される。本例では、通信チャネル101を用いてこの顧客から音声コマンドまたは発話をコンバーサントIVRシステム18に提供しさらにビデオ画像と音声の情報をコンバーサントIVRシステム18からこの顧客に提供する。この顧客は音声コマンドを電話機器30を経て信号バス102でボイスビュー（Voice View）モデム104に入力する（VoiceViewはRadish通信システムズ社の登録商標である）（ボイスビューモデムはRadish通信システムズ社から提供され利用できる）。

【0022】パーソナル・コンピュータ98からデジタル・データはモデム104に信号バス106で提供される。（本例ではモデム104はデジタル情報を信号バス106で受信しない。しかしクレジット・カードの画像や口座番号を表すデータのようなデジタル・データはモデム104にコンピュータ98と信号バス106を介して提供できる。）モデム104はこのユーザの発話を信号バス108で情報通信ネットワーク10を経て信号バス110に送信する。ボイスビュー モデム112は情報通信ネットワーク10から信号バス110でこの信号を受信しこれをアナログとデジタルの出力に分割する。アナログ出力は信号バス114でコンバーサントIVRシステム18に提供される。

【0023】またデジタル出力は信号バス116でコンバーサントIVRシステム18に提供される。（前述のように本例ではモデム104からモデム112に進むデジタル情報はない。しかし前述のようにPC98はクレジットカード情報のような情報をモデム104、112を用いてコンバーサントIVRシステム18に伝えることができる。）コンバーサントIVRシステム18は電話機器30におけるこのユーザへその音声メニューを信号バス114とモデム112を介し通信チャネル110を通りモデム104と信号バス102を経て伝える。次にこのユーザはこのメニューから選択するため音声コマンドを入力する。この音声コマンドはコンバーサントIVRシステム18によって翻訳されこれを用いてこのユーザに追加メニューとその選択画像を提供する。

【0024】先に図1に関して述べたように、選択画像に音声情報を含むことができる。さらに一連のビデオ画像を提供しこの顧客が動画ビデオを受信できる。コンバーサントIVRシステム18はこの顧客が選択したビデオ画像データおよび／または音声データを提供するがこれはコンバーサントIVRシステム18のメモリ内また

は記憶装置60内に蓄積した音声ファイルおよび／またはビデオ画像ファイルをアクセスして行われる。この音声データはアナログ形で信号バス114を介しモデム112に提供され、このビデオ画像データは信号116を介しモデム112に提供される。前述のようにこの画像データはそのJPEGフォーマットのようないずれかの通常使用されるフォーマットとすることができる。

【0025】モデム112はこの音声情報を信号バス114からモデム104に送る。しかし画像データが信号バス116で利用されていると、この画像データが送られその音声情報の伝送はこの画像データの伝送が終了するまでストップする。モデム104はこの音声と画像のデータを分離しその音声信号を信号バス102で電話機器30にこの顧客の聴取用に提供する。モデム104はこのビデオ画像データを信号バス106でパーソナル・コンピュータ98に提供する。パーソナル・コンピュータ98はモデム104から受信したその標準のビデオ画像データ・ファイルを用いてこの送信ビデオ画像をディスプレイまたはモニタ100に生成する。

【0026】ここで注記する必要があることは、この顧客にJPEGタイプの画像データを送ることに加えて、さらにまたファクシミリ画像データをこの顧客に信号バス116で送ることができる。ファクシミリ画像データを信号バス116で送る場合、モデム104はこの情報をPC98へ信号バス106を通して提供する。次にPC98はこのファクシミリ画像データをディスプレイ100に表示するかまたはこのファクシミリ画像をプリンタを用いてプリントする。以上の説明は、本発明の実施の二三の形態例に関するもので、この技術分野の当業者であれば、本発明の種々の変形例が考え得るが、それらはいずれも本発明の技術的範囲に包含される。尚、特許請求の範囲に記載した参照番号は発明の容易なる理解のために、その技術的範囲を制限するよう解釈されるべきではない。

#### 【0027】

【発明の効果】以上述べたごとく、本発明の方法によって、例えば、顧客は情報通信ネットワークでこの顧客に関心のある商品画像を選択しさらにこの顧客の都合に応じてショッピングしさらに音声コマンドを用いて提供商品画像を制御し購入商品を選択することができ有効である。

#### 【図面の簡単な説明】

【図1】情報通信ネットワークを通る複数のチャネルを用いて顧客に画像データを提供するよう情報通信ネットワークを利用する実施の形態例を示す図である。

【図2】顧客に画像データを提供するようこの情報通信ネットワークを通る単一チャネルを利用する実施の形態例を示す図である。

【図3】顧客に画像データを提供するようこの情報通信ネットワークを通る単一チャネルを利用する別の実施の

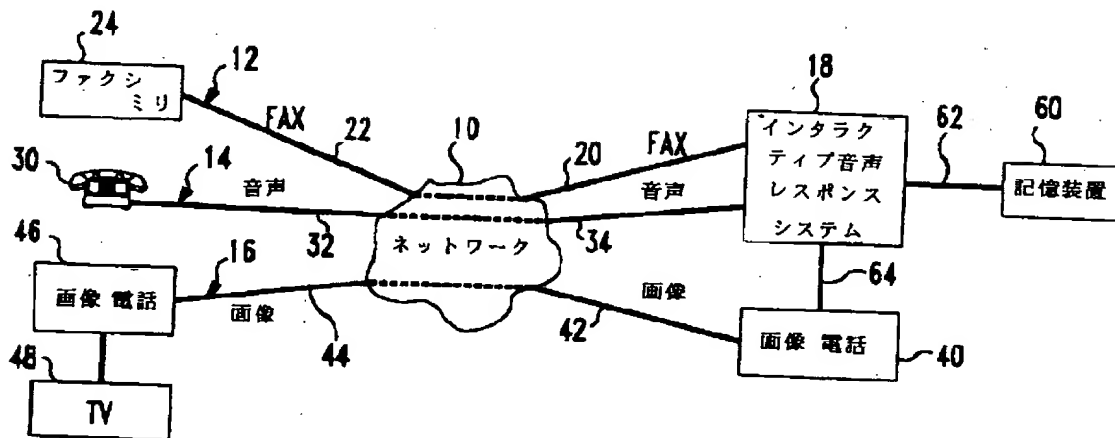
形態例を示す図である。

【符号の説明】

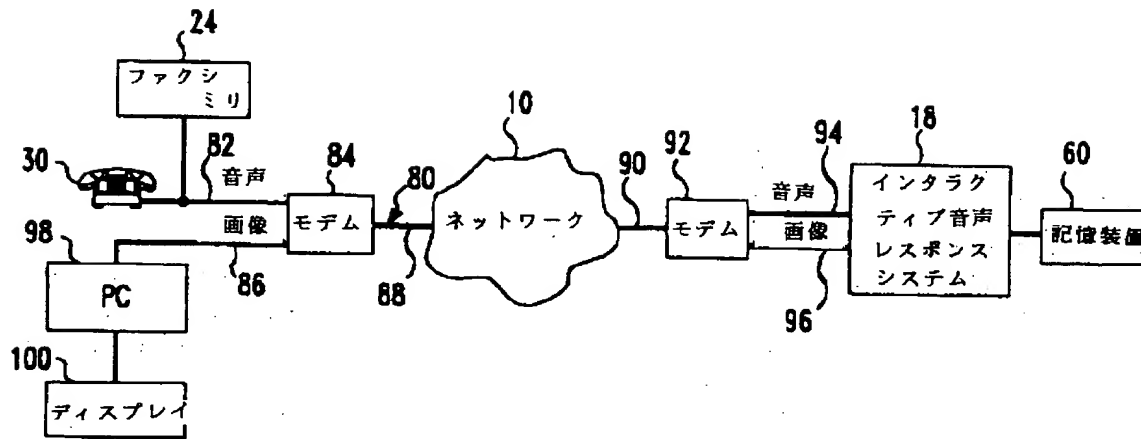
- 10 情報通信ネットワーク  
 12 通信チャンネル  
 14 通信チャンネル  
 16 通信チャンネル  
 18 (音声認識装置) インタラクティブ音声レスポンス (IVR) システム  
 20 信号バス  
 22 信号バス  
 24 ファクシミリ端末  
 30 電話機器  
 32 信号バス  
 34 信号バス  
 40 ビデオ電話  
 42 信号バス  
 44 信号バス  
 46 ビデオ電話  
 48 テレビジョン  
 60 記憶装置  
 62 信号バス

- 64 信号バス  
 80 通信チャンネル  
 82 信号バス  
 84 モデム  
 86 信号バス  
 88 信号バス  
 90 信号バス  
 92 モデム  
 94 信号バス  
 10 96 信号バス  
 98 パーソナル・コンピュータ  
 100 ディスプレイまたはビデオ・モニタ  
 101 通信チャンネル  
 102 信号バス  
 104 モデム  
 106 信号バス  
 110 通信チャンネル  
 112 モデム  
 114 信号バス  
 20 116 信号バス

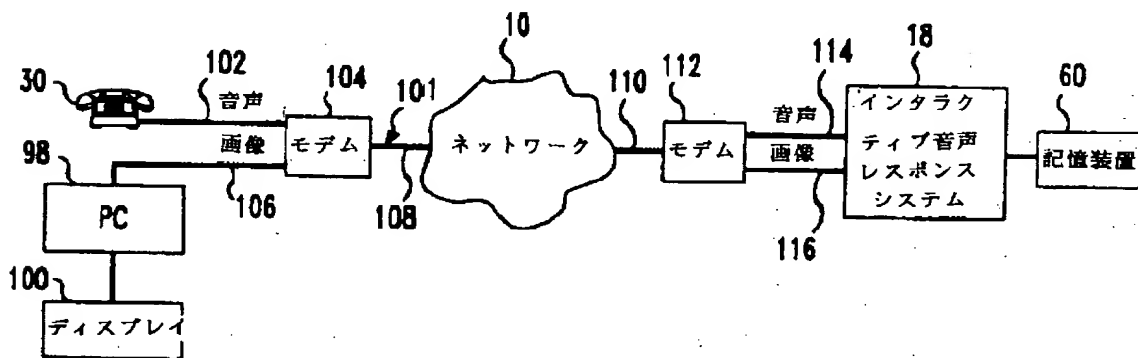
【図1】



【図 2】



【図 3】



フロントページの続き

(72)発明者 マーク エー. ボルスター  
 アメリカ合衆国, 43221 オハイオ, コロ  
 ンブス, ウォーターグラス ヒル ドライ  
 ブ 3510